Center for Independent Experts (CIE) Reviewer's Independent Report on the 2016 CIE Review of assessments of Pacific cod stocks in the Eastern Bering Sea and Aleutian Islands

Prepared by

Neil Klaer

Prepared for

Center for Independent Experts

Review Meeting
Alaska Fisheries Science Center
Seattle, Washington
16-19 February 2016

Contents

Executive Summary

The 2016 CIE Review of assessments of Pacific cod (*Gadus macrocephalus*) stocks in the Eastern Bering Sea (EBS) and Aleutian Islands (AI) met in Seattle, Washington, from Tuesday to Friday, 16-19 February 2016. The meeting was chaired by Anne Hollowed from the Alaska Fisheries Science Centre. The review panel was composed of Robin Cook, Jean-Jacques Maguire, and Neil Klaer from the Center for Independent Experts (CIE). The meeting generally followed the draft agenda and included presentations by the stock assessment team (STAT) mixed with questions and open discussion. Additional analyses were requested by the Panel from the STAT and the results of those were also subsequently presented. After input data and model presentations and general discussions, the remaining days of the meeting were devoted to the examination of various aspects of the models through a request and response process. The structure of the meeting was to concentrate on providing answers to the specific questions in the terms of reference (ToRs), rather than arriving at a consensus view on any particular model. Each member of the Panel has written his own independent report on the meeting.

Findings for Pacific cod

Generally, if a data source provides useful information, can fit within an assessment model structure, has been shown to be reliably collected and standardized, and is likely to be unbiased or bias can be accounted for, then it should be included in the stock assessment model. I have recommended that the IPHC and NMFS longline surveys and associated composition data be included in both the EBS and AI models conditional on documentation that examines those surveys for potential bias regarding Pacific cod. The IPHC in the EBS in particular may require truncation to eliminate suspect point estimates.

Input sample sizes for composition data have an influence on assessment results and it has also become generally accepted practice for those sample sizes to more reflect the number of sampled fishing trips.

For relative weighting of various data sets, I recommend estimation of an additional sd for all abundance indices, and either the Francis or harmonic mean weighting procedures for composition data.

Options for selectivity patterns are primarily among simple logistic and double-normal by size, and random walk by age. The simplest pattern that allows reasonable model fit to available composition data should be used. The most complex random walk by age pattern is most suited for application to combined fisheries composed of differing gear types, although there may be a question about the implementation of it in SS regarding large final gradients.

Allowing time-varying selectivity that is a random walk by age annually for a fishery with multiple gear types is an innovation that I have not seen previously. As many current SS assessments grapple with highly partitioned fishery data, such a procedure has the potential for wide application. I am reluctant to agree on its use without a supporting simulation study that confirms its equivalence or even superiority to a high degree of data partitioning. Such a study would be reasonably easy to design and carry out. However, I am willing to agree that it seems to provide a good resolution to the problem for the fishery selectivity in the EBS models.

Time variability should be allowed in a parameter when there is an available reliable data source that fairly directly measures such a change, and that a trend exists in that data source that needs to be captured by the assessment model. This situation only currently appears to exist for recruitment and fishery selectivity in the EBS model.

Models examined during the review for the EBS seem to fairly clearly demonstrate that the trawl survey selectivity is dome-shaped. However, the possibility that the survey selectivity is in fact asymptotic has not been eliminated. The extent of the survey dome-shape may, for example, be confounded with M. It may be that different data sources are in conflict about the estimated value for M that can be diagnosed with a Piner profile plot of likelihood components. Exploration of age-specific M (e.g. starting with a Lorenzen function) could also be done. A range of plausible alternative models should be explored, and the extent of the estimated dome selectivity for the trawl survey examined for each to see if the dome is consistently required. However, as the extent of the trawl survey dome is probably one of the major axes of uncertainty in the EBS model at present, it should remain freely estimated and informed by the available data in any chosen base model, possibly with forcing more or less dome as sensitivity analyses in the final assessment.

Models that estimate the shelf bottom trawl survey q using a fairly non-informative prior (as in EBS model 15.6) should currently be preferred. Agreed bounds on prior survey q point estimates can be used as one of the acceptance criteria for particular models. I personally have a fairly high tolerance for those values (based however, on only a limited background knowledge for this particular survey), and am comfortable with at least a factor of 2.0 (0.5 – 2.0 times the initial point estimates). Should additional surveys be added to the models, q values for Pacific cod for those are less well understood, and non-restrictive priors for those are preferable, with q estimated. Work should be commenced on the development of a prior distribution for EBS shelf bottom trawl survey q that can be generally agreed.

I have made additional comments not specific to the ToRs on diagnosis and potential correction of retrospective patterns, catch uncertainty, alternative values for steepness, regime change and inclusion of marginal age composition vs conditional age-at-length data.

1 Introduction

1.1 Background

The 2016 CIE Review of assessments of Pacific cod (*Gadus macrocephalus*) stocks in the Eastern Bering Sea (EBS) and Aleutian Islands (AI) met in Seattle, Washington, from Tuesday to Friday, 16-19 February 2016. The meeting was chaired by Anne Hollowed from the Alaska Fisheries Science Centre. The review panel (the Panel) was composed of Robin Cook, Jean-Jacques Maguire and Neil Klaer from the Center for Independent Experts (CIE).

Draft stock assessment reports, as well as associated background documents, were made available via a Google drive and subsequently from a public website to the Panel on 2 February prior to the review meeting. During the meeting, all documents were available electronically via the same website, and meeting presentations and additional documents and model runs made during and after the meeting were also posted there.

The meeting generally followed the draft agenda and included presentations by the stock assessment team (STAT) mixed with questions and open discussion. Additional analyses were requested by the Panel from the STAT and the results of those were also subsequently presented.

1.2 Review Activities

After input data and model presentations and general discussions, the remaining days of the meeting were devoted to the examination of various aspects of the models through a request and response process. Requests were generally devised to allow the terms of reference (ToRs) to be addressed, although some additional ones were also made. There was not an expectation to produce an agreed meeting summary report, so this review differed from most that I have attended. This meant that a consensus view on particular base models was not required, allowing the meeting to better concentrate on providing answers to the specific questions in the ToRs. The only real pressure during the meeting was to ensure that requests had covered issues that the reviewers deemed necessary for addressing the ToRs. A fairly large list (Table 1) was constructed that was not possible to fully complete during the meeting, so requests were prioritized. All of the requests except low priority ones were completed by the middle of the following week. Many thanks to Grant Thompson (senior assessment author) and Steve Barbeaux (support assessment author) in completing those requests.

Each member of the Panel has written their own report following the meeting without collaboration, so it should be very interesting to compare and contrast those for common and differing opinions.

Table 1 Exploratory Run Requests

Green=completed during meeting; yellow high or medium priority; no color useful but not critical. Model 15.7 = authors best AI model, Model 16.1 = Model 15.7 with simpler selectivity; Model 15.6 = author recommended, and CIE's selected model, for EBS.

AI

1. H: Explore SS Truncated time series 1997: Modified Model 3 + either one of M15.4
2. M: Lambda on age and length comp stepped down to explore role of mean sample size (5 years of survey age composition) (Long – term)
3. H: AI 15.7 explore alternative in simpler fishery selectivity function (e.g., Time invariant – double normal), run 3 retrospective.
4. Try reducing sigma on annual devs of selectivity or selecting time blocks to reflect shifts in the relative ratio of trawl to longline in fishery
5. Start whole model in 1991 (was previously done in 2014) and Turn off time varying selectivity for the trawl survey (reference as 16.1)
6. Add NMFS LL survey to 15.7 and 16.1
7. Add IPHC LL survey to 15.7 and 16.1
8. Add both LL surveys to 15.7 and 16.1
9. Individual estimation of age selectivity for survey and fishery

EBS Rationale for selecting Model 15.6 Good retrospective, rational changes to model.

1. Update Model 15.6 with 2015 data, (Q = 0.88 with new data added);
2. M: Make run Model 15.6 with mean size-at-age turn on
3. H: Ratio - Catch curve estimate Z
4. H: Index time series IPHC, NMFS Longline, add slope survey
5. H: Plot Size comps of IPHC, NMFS longline, and slope survey
6. H: Adding IPHC to one or two models 11.5, Model 15.6
7. L: Explore inclusion of fishery age composition 4 years available (08, 09, 10, 11)
8. M: Try steepness run with 0.7 for 11.5 and 15.6
9. M: Model 15.6 (shelf survey only) Current sigma, Turn off time varying Q; half of sigma for Q; AIC or DIC? Weighing model parsimony vs improved realism.
10. H: Changing catch CVs (Fishing Mortality weights) Model 15.6
11. M: Piner profile M and R0, no more than 10 points
12. L: Francis weighting on current models,
13. L: Number of stations and sample size for ll surveys and slope.
14. Explore taking out time varying selectivity in 15.6
15. Retrospective runs of 15.6
16. H: run with EBS slope survey only
17. H:Turn on SD estimation for everything except the shelf trawl survey
18. H: run with EBS slope, IPHC, NMFS LL

2 Stock assessments

2.1 Terms of reference

The Panel considered the assessments in light of the terms of reference provided as follows:

1. Evaluate and provide recommendations on data used in the assessment models. In particular:

   a. Should data from the IPHC longline survey be used in either assessment?

   b. Should data from the NMFS longline survey be used in either assessment?

2. Evaluate and provide recommendations on model structure, assumptions, and estimation procedures. In particular:

   a. How should the various data sets be weighted?

   b. What form (i.e., Stock Synthesis "pattern") should be used for the selectivity functions?

   c. Should the models be structured with respect to season?

   d. Should the models be structured with respect to gear type?

   e. How much time variability should be allowed, and in which parameters?

   f. What constraints, if any, should be placed on survey selectivity at older ages?

   g. What constraints, if any, should be placed on survey catchability?

   h. How should large gradients be dealt with in otherwise apparently converged models?

   i. Anything else on which the reviewers care to comment.

2.2 Findings by term of reference

The comments below refer to aspects that were discussed during the review, but include my own additional commentary for preparation of this CIE report.

2.2.1 Evaluate and provide recommendations on data used in the assessment models.

As a general principle, we all understand that data to be potentially included in a stock assessment model first need an examination to determine whether they measure important aspects of stock dynamics that can be included in a stock assessment model, are collected and standardized in a rigorous manner, and are likely to be unbiased or any bias has been measured and can be accounted for. Ideally, this examination for each separate input data set would be well documented, updated as required, and provided as support information for any stock assessments. Most stock assessments do not reach this ideal. For the EBS and AI assessments, such data documentation specifically for stock assessment support does not exist. However, during the review, presentations were made that described data collection methodologies and the process used to prepare the data for use in stock assessments which could form the basis for such documentation.

The most difficult input data question regards possible bias. Normally, it is the data collectors who have the most information about changes in collection procedures, unexpected changes in data signals, potential for non-representative sampling and the like. Input data documentation should include accounts by the data collectors on these aspects, and the potential bias that may have been introduced. Where several data sources provide similar information (e.g. alternative survey abundance indices with similar gear selection), it may also be useful to ask data collectors to rank the alternatives according to potential bias. Such information may then be used by stock assessment authors when preferentially weighting various data sets.

A particular example examined during the review that illustrates the usefulness of improved documentation was for the AI trawl survey abundance index and associated composition data. A list of 10 historical changes in survey design was provided, but it was acknowledged that the input data had not been subjected to a detailed examination regarding those changes to potentially quantify their effects. As some changes appeared to be substantial but also open to desk-top investigation (e.g. any apparent shift in selectivity pattern due to the change from 30 to 15 minute trawls), my initial reaction was to not use the series trend until appropriate investigations had been made. Subsequent discussions concluded, with the help of data collectors, that changes since 1997 in survey methodology were unlikely to have caused substantial bias in the index, so it was agreed that the index was usable from that year forward.

Generally, if a data source provides useful information, can fit within an assessment model structure, has been shown to be reliably collected and standardized, and is likely to be unbiased or bias can be accounted for, then it should be included in the stock assessment model.

In particular:

2.2.1a Should data from the IPHC longline survey be used in either assessment?

EBS

The shelf trawl survey in most/all EBS models appears to require dome selectivity in comparison with the fisheries regardless of whether the fisheries are highly partitioned according to gear and season, or selectivity is allowed to change through time (e.g. both models 11.5 and 15.6). Ideally, abundance index size/age selection would be reflective of the population – i.e. asymptotic at a low age. Models that include the trawl survey alone have considerable flexibility to alter abundance trends for older age classes not well indexed that may have a heavy influence on population SSB trends. There is an advantage therefore, to include index information for those older age classes if such indices exist. In this case, candidates are the IPHC and NMFS longline and the slope trawl surveys.

In all cases (and IPHC in particular), the available additional surveys were primarily designed to index species other than Pacific cod. Desk-top studies of the suitability of application of these surveys as potential indices of abundance to Pacific cod in particular are currently unavailable, so judgment of whether to include them into an active assessment model is only evaluated here based on presentations of survey procedures during the review, general comparisons among available indices, and the apparent performance of models that include various index combinations.

The IPHC primary objectives are to provide CPUE, length and age composition, information on abundance distributional changes for juveniles and adults for Pacific halibut. Secondary objectives are to provide information on bycatch species and a platform for specialized projects. We learned through presentations that a number of factors (different hook size to commercial Pacific cod fishing, first 20 hooks per skate sampled for bycatch, bait used, areas sampled) may not be optimal for Pacific cod, but Pacific cod are the most-often encountered bycatch species by the survey (at least in Areas 4A, 4B, 4C and 4D – covering the EBS and AI regions). This suggests that IPHC survey trends at least require examination, and that there are no reasons yet identified that imply an index bias, just sources of possibly random measurement error.

The aggregated size composition from the IPHC survey indicates a selected size range well to the right of the shelf trawl survey in the EBS, and slightly to the right of the longline fishery, NMFS longline survey and slope trawl survey (Figure 1). This indicates that the IPHC index can potentially provide useful abundance information for the older age-classes that are not indexed by the shelf trawl survey if that survey selectivity is dome-shaped.

A comparison of general index trends in the EBS (Figure 2) does not show a lot of consistency among available indices, although the different selectivity associated with those indices makes interpretation more difficult. The IPHC survey seems to exhibit trends that are least consistent with the other available indices. A shift of the IPHC survey several years to the left shows perhaps some consistency with the trawl survey. Biologically, it is not possible for the true abundance of older year classes in the Pacific cod population to change radically from one year to the next. There are two substantial drops in the IPHC index that seem biologically implausible – in 1999 and 2005. Further work is needed to investigate the cause of these changes in particular, and whether the index requires refinement in application to Pacific cod.

Among the meeting requests were those that included various new index combinations to be added to the EBS model, while also estimating an additional sd. The additional sd accounts for apparent error that is required to be added to an index for the model to be balanced, given the information from all other data sources in the model (model 15.6 extra sd). That model adds a large sd value to the IPHC index, mostly to better account for the apparent error in the 1999 index value.

Before deciding to include the IPHC longline index and associated lengths in a proposed central EBS SS model, an investigation into the properties of the EBS IPHC longline index in relation to Pacific cod in particular should be done. The investigation should examine the 1999 and 2005 points especially to see if justification exists for exclusion – perhaps by starting the IPHC index in 2000. If the resulting index is found unlikely to be biased, then I recommend inclusion in the model with additional sd estimated.

AI

Most of the effort of the meeting was directed towards investigation of the properties of the EBS assessment model, as an SS assessment is already the agreed approach for that region. The AI is currently a Tier 5 that essentially applies a smoother through trawl survey estimates of total biomass. However, the assessed trend in biomass is less important than the most recent estimate in the provision of management advice. It was hoped that if reasonable approaches to data and modeling can be determined for the EBS, then many of those same approaches could also be applied to the AI region. My initial thought was that an agreed EBS model could be entirely transferred to the AI, but it was shown during the meeting that simplification of the AI model can lead to improved model behavior – particularly regarding retrospective patterns. Indeed, the removal of time-varying factors can sometimes improve retrospective behavior, possibly in conflict with general conclusions of recent publications (e.g. "when retrospective patterns are observed in a stock assessment, they are often corrected by introducing estimation of a time-varying parameter (usually selectivity, M or q)", Hurtado-Ferro et al. 2014).

The AI model is the same as for the EBS in that the trawl survey selectivity appears to be domed and to the left of the fishery, and that the IPHC survey has potential use for providing an index for older age-classes (Figure 3). Even without estimation of an additional sd, the IPHC index can be reasonably well fitted by the model, with 2012 being the largest influential residual. Further work on choice of a more appropriate selectivity function other than double-normal (or by changing the freedom of certain double-normal parameters) would probably improve the overall fit to IPHC lengths (Figure 4).

Before deciding to include the IPHC longline index and associated lengths in a proposed central AI SS model, an investigation into the properties of the AI IPHC longline index in relation to Pacific cod in particular should be done. If the index is found unlikely to be biased, then I recommend inclusion in the model with additional sd estimated.

Figure 1 EBS longterm size composition comparison (indices, longline fishery)

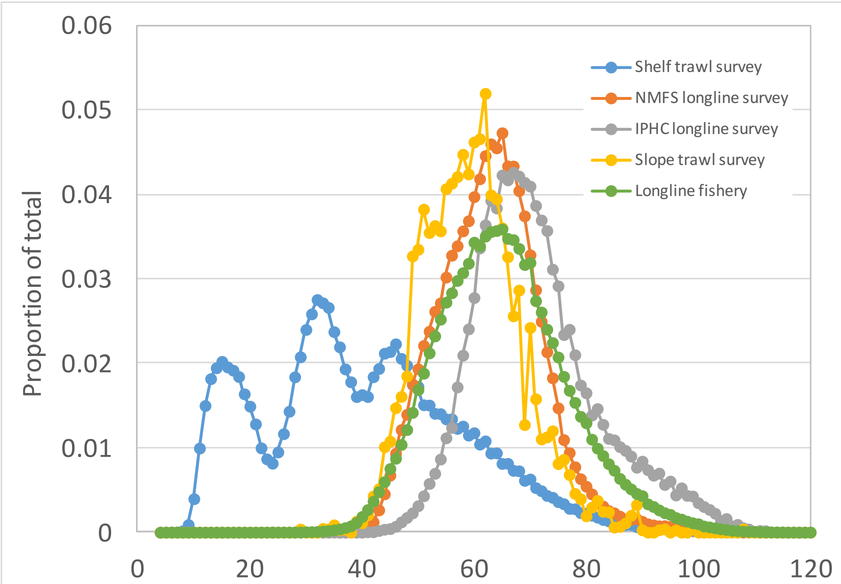


Figure 2 EBS survey index comparison

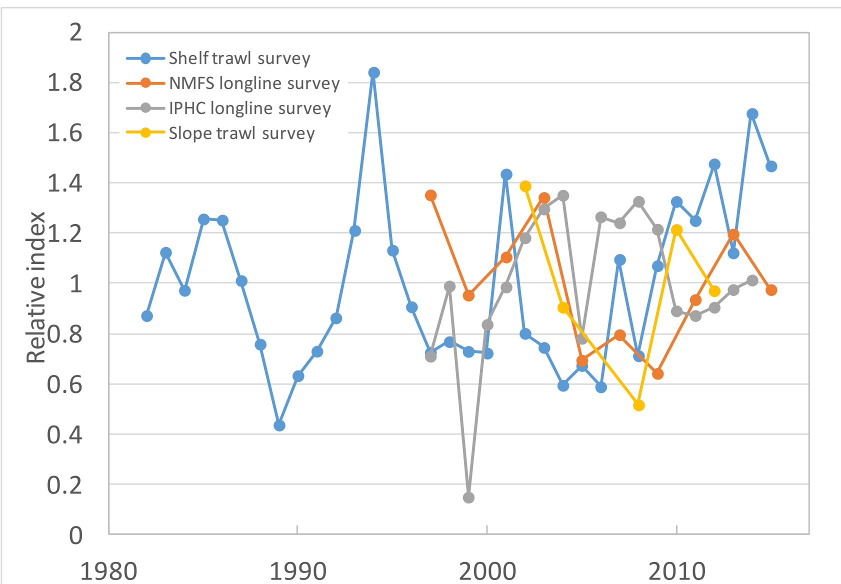


Figure 3 AI length composition summary – model survey index comparison AI model 16.6 (16.1 plus IPHC)

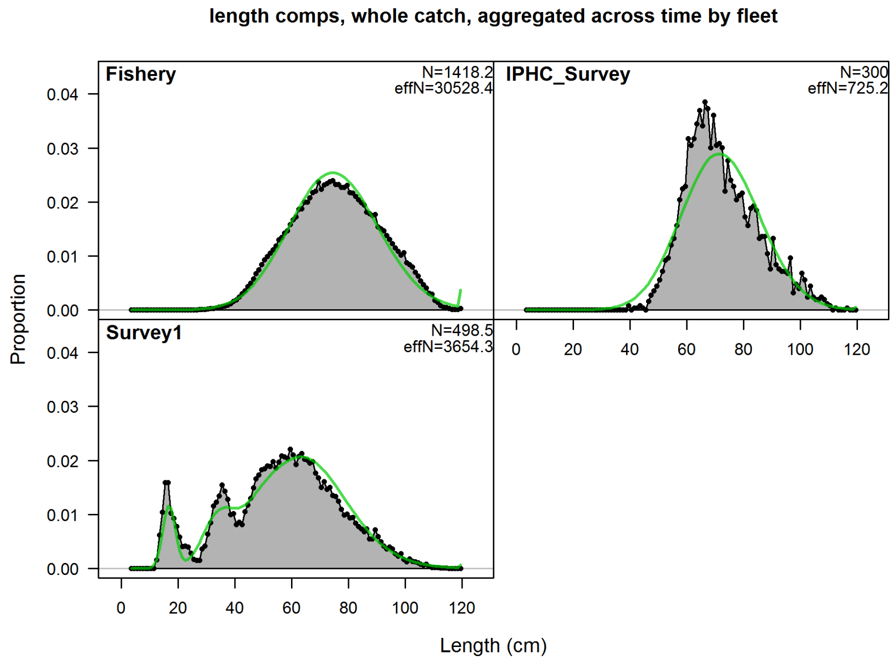
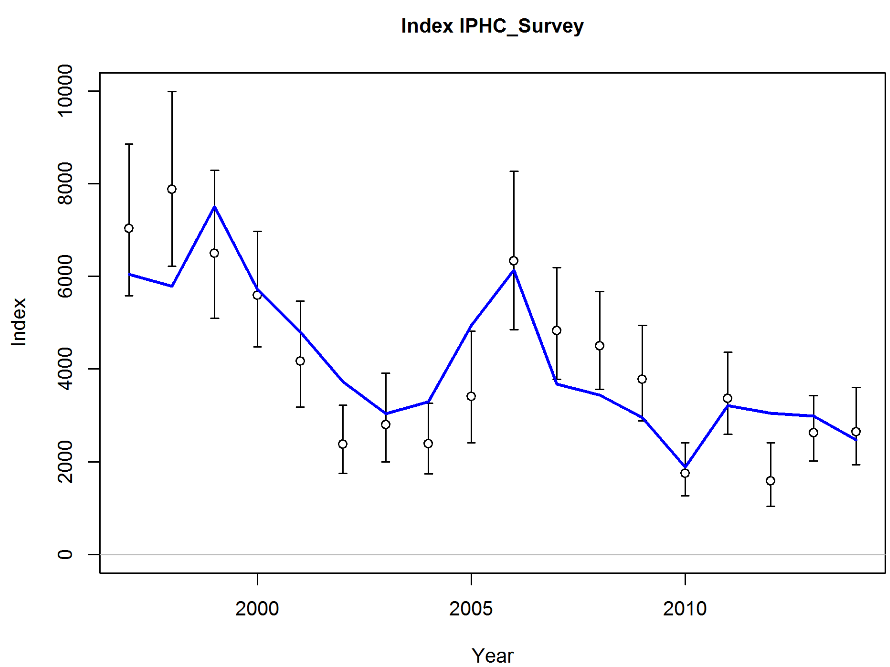


Figure 4 AI IPHC index fit for AI model 16.6 (16.1 plus IPHC)



2.2.1b Should data from the NMFS longline survey be used in either assessment?

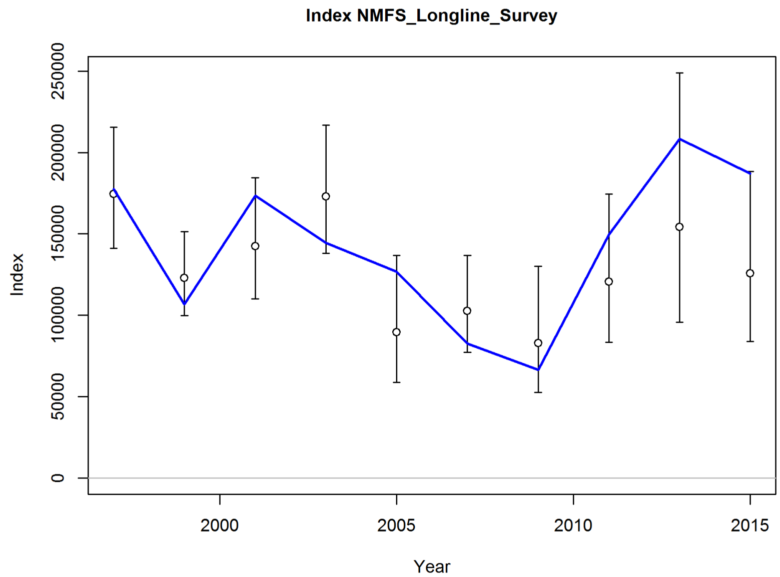
The primary aim of the NMFS longline survey is to collect abundance, composition and bycatch information for Sablefish. Again, a desktop study has not been made to determine whether the survey is potentially biased with respect to Pacific cod abundance. Indices for Pacific cod are available for EBS and AI, although the survey does not cover the western AI region. Age compositions are not collected for Pacific cod by this survey, but there are many lengths collected.

During the review a question was raised about the possible over-weighting of surveys, particularly through the use of multiple longline surveys in a single model, and it was suggested that they could potentially be combined before addition to the model. My own preference on this is to keep independent data sources separate, and to let additional sd estimation weight each based on goodness of fit with all other data sources in the model. I think it is an advantage if independently collected indices show similar trends for the same size/age classes in the population, and should therefore receive more weight in those circumstances. Alternatively, conflicting indices should be down-weighted in an objective manner.

EBS

Aggregated lengths for the EBS show that the NMFS longline survey seems to catch about the same size fish as the longline fishery, but not as many of the very largest fish as does either the fishery or IPHC surveys (Figure 1). Relative index trends show that the NMFS longline seems potentially more consistent with the shelf survey than the IPHC survey if shifted several years to the left (Figure 2). The NMFS longline survey does not show large changes in abundance that are biologically implausible as the IPHC survey does. Addition of the index to the model even without additional sd estimation shows a reasonable fit by the model (Figure 5). Of potential stock concern is that the NMFS longline survey is generally under the expected survey abundance since 2010 (Figure 5), suggesting that information on larger fish in the population added by this survey leads to a more pessimistic assessment of overall stock depletion (as indeed shown by model 15.6A results). However, the model is not fully tuned, so such supposition may be premature. However, it does highlight that if the index is to be used, some evaluation of possible bias in relation to Pacific cod, perhaps most importantly since 2010 is required. The model that includes the NMFS longline survey is able to fit the associated length compositions well.

Figure 5 NMFS longline survey index fit for EBS model 15.6A (15.6 plus NMFS longline survey)



Before deciding to include the NMFS longline index and associated lengths in a proposed central EBS SS model, an investigation into the properties of the EBS NMFS longline index in relation to Pacific cod in particular should be done. The investigation should particularly examine possible bias in the index since 2010 as this appears to be influential on assessment results. If the index is found unlikely to be biased, then I recommend inclusion in the model with additional sd estimated.

AI

The overall fits by the AI model to lengths (Figure 6) and the abundance index appear reasonable. Abundance index point estimates for 2004 and 2014 appear to most conflict with other information in the AI model.

Before deciding to include the NMFS longline index and associated lengths in a proposed central AI SS model, an investigation into the properties of the AI NMFS longline index in relation to Pacific cod in particular should be done. If the index is found unlikely to be biased, then I recommend inclusion in the model with additional sd estimated.

Figure 6 AI length composition summary – model survey index comparison AI model 16.4 (16.1 plus NMFS longline)

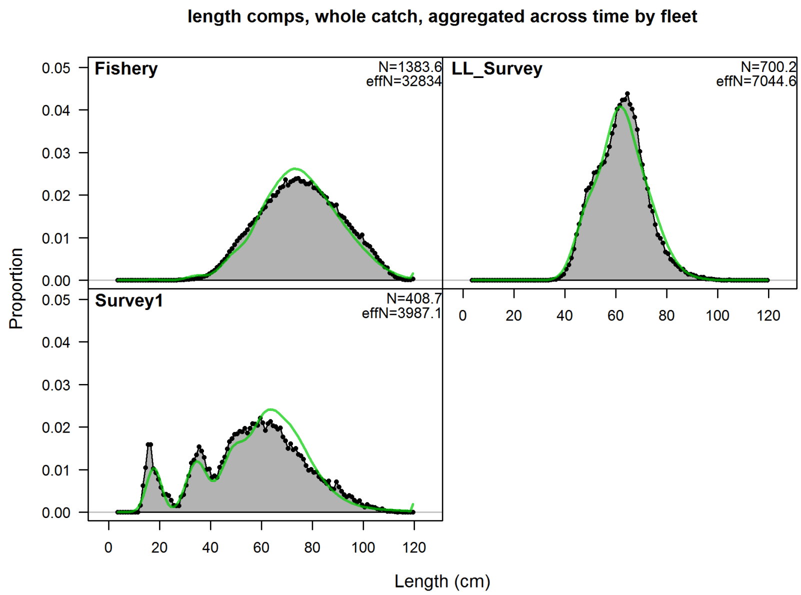
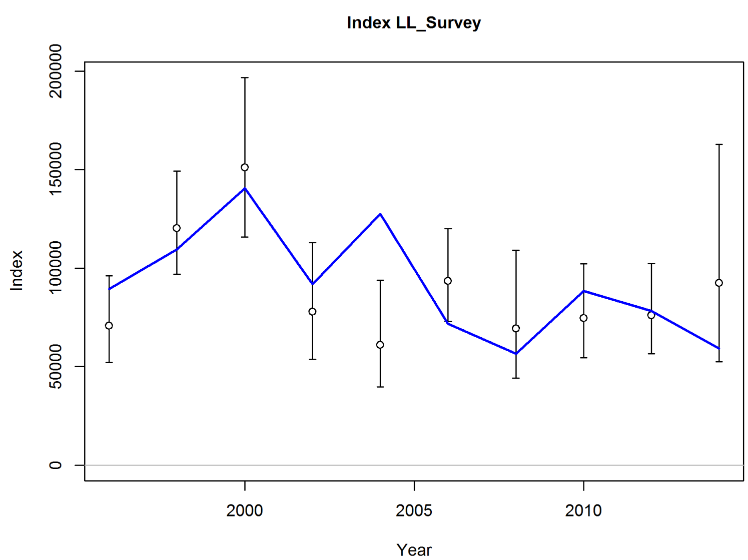


Figure 7 AI NMFS longline index fit for AI model 16.4 (16.1 plus NMFS longline)



2.2.2 Evaluate and provide recommendations on model structure, assumptions, and estimation procedures.

In particular:

2.2.2a *How should the various data sets be weighted?*

For abundance index data, iterative reweighting to potentially allow additional index error was previously an accepted procedure for many US and Australian stock synthesis assessments. Such iteration was done manually, and more recently the ability to internally estimate additional index error (via an additional sd) has been added as an option to SS. Use of that option has become accepted practice for many recent assessments. Estimation of additional index error is normally done for all indices included in a stock assessment as (perhaps in my naive interpretation), the input variability usually only accounts for measurement error and the process error component is unknown.

Input sample sizes for composition data have an influence on assessment results and it has also become generally accepted practice for those sample sizes to more reflect the number of sampled fishing trips, rather than the number of fish measured.

Relative data weighting in stock assessments for composition data and the goal of standardized approaches has been the subject of recent and ongoing research particularly in the US west-coast, and the subject of a Center for the Advancement of Population Assessment Methodology (CAPAM) workshop in La Jolla, CA in October of 2015 (http://www.capamresearch.org/data-weighting/workshop). While there has been some recent narrowing down of agreed procedures among US west-coast stock assessors, it has also been recognized that it is not currently possible to recommend default procedures for composition and conditional age-at-length (CAAL) data. There is agreement that the Francis weighting approach is more appropriate in cases where the model is not correctly specified as it takes autocorrelation among composition data into account. It is also agreed that for a correctly specified model, the McAllister-Ianelli harmonic mean weighting method works well. Both of these procedures have been extended from marginal length or age composition data to conditional age-at-length (Francis A and B methods are available for CAAL, with Francis B potentially preferred). A possible further development that may provide a direction forward is using the Dirichlet multinomial likelihood (Thorson, 2014), although this method will require review and implementation in SS before it may be used. Recent simulation work has shown that the McAllister-Ianelli arithmetic mean procedure is inferior to other methods (Punt, *In press*).

2.2.2b *What form (i.e., Stock Synthesis "pattern") should be used for the selectivity functions?*

SS provides a large number of selectivity pattern options (14 size and 12 age patterns excluding special, discontinued and mirror – SS user manual v 3.24s). By far the most commonly used patterns in recent stock assessments are logistic for simple asymptotic selectivity or the double-normal (most often size pattern 24 or age pattern 20) where selectivity is allowed to be dome-shaped. The flexibility of the double-normal is usually sufficient to account for the wide range of single-peaked shapes that may be expected from a single fishing gear type. It is also possible to combine size and age selectivity patterns for a fishery or survey and to have differential selectivity by sex to, for example, account for reduced availability of older females in the population. To most easily account for "odd-shaped" selection that may be due to, for example, a combined fishery composed of several gear types, SS provides an age based selection pattern that generates an age-based random walk (age pattern 17).

Normally, fishery and survey selection is assumed to be primarily a length-based process as fishing gear selection is usually size-dependent. However, selectivity in an assessment model combines gear vulnerability with availability. Whether availability (e.g. due to migration, aggregation [e.g. for spawning], schooling) is age- or length-based is a more difficult question, so although length-based selection may be preferred for modeling, a case can still be made for age-based selectivity.

Generally, the selectivity pattern should be chosen (most likely from the options above) that has the fewest parameters, and allows an acceptable fit to the available composition data (e.g. no bands at particular lengths of significant length composition residuals). As surveys are designed to at least use the same fishing gear throughout, a good reason to use more complex patterns than logistic or double-normal would be required for those. If a fishery has fairly homogenous gear, a similar argument applies there as well. In the case of a fishery with mixed gear types, an opportunity exists to use a less restricted pattern shape, as provided by the age-based random walk. At present, I don't think a random-walk length-based pattern is available, so selectivity in that case is restricted to being age-based.

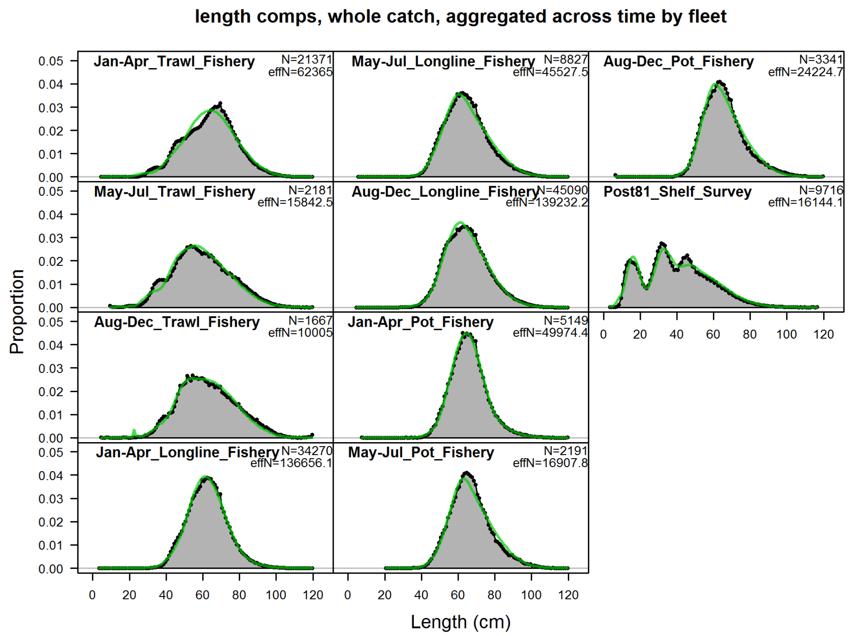
2.2.2c *Should the models be structured with respect to season?*

It is usual practice for SS models to separate input data from surveys and fisheries that have demonstrably different selectivity if data are available to do so. Normally, the minimum requirement to allow data partitioning according to season, gear type or area is that a number of years of length or age composition data that are believed to be representatively sampled are available within each partition. Partitioning of composition data is only usually necessary if summary length/age compositions from comparable partitions show obvious apparent selectivity differences. Partitioning may also be required for abundance indices if different trends are observed by partition.

Models that specifically address the exploration of alternative structures regarding selectivity partitions have been developed and were presented for the EBS, so the discussion here will be confined to models from that region.

Simple examination of aggregated length data for the EBS shelf trawl survey, the slope survey, longline fishery and NMFS and IPHC longline surveys (Figure 1) show a marked difference in the shelf trawl survey to all of the others. Unfortunately, the trawl and pot fisheries were not included, but we know from diagnostic output from model 11.5 that trawl fishery selectivity seems to be intermediate between the trawl survey and longline fishery, and the pot fishery seems similar to the longline fishery (Figure 8). Also notable is that the Jan-Apr trawl fishery lengths show a peak that is consistent with longline fisheries during that period only, which corresponds to the spawning season. Conjecture has been made about possible movement of larger fish from the NBS area, although another explanation may be the movement of larger fish from waters targeted by the longline and pot fisheries into shelf trawl areas during the spawning season. There is little information available from tagging and none that can address the question of movement in and out of the NBS. The shelf trawl survey is made outside of the spawning season, and at that time, less of the larger fish seem to be available on the shelf, although tagging of a small number of fish does indicate apparent random movement of fish over the shelf during that time.

Figure 8 EBS aggregated length compositions and fit for model 11.5



Fishery selection

For modeling purposes, the model only requires that the composition of the fishery catches be adequately accounted for each year, and the more important population abundance trends are taken from surveys (at least for the models here). The difference in trawl fishery selection by season seems to be a feature that can be addressed through seasonal model structure. This is done to some extent with model 11.5, but the fit to the Jan-Apr trawl fishery length composition by that model is not particularly good (Figure 8). In addition to gear/season partitioning, a large number of time blocks that allow selectivity to vary through time have been used in model 11.5. It may be questioned whether such fine scale partitioning of the data are supportable if partitioning and blocking first needs to be justified depending on whether prior data examination or independent knowledge about changes in practices suggests that all of those partitions are necessary, and that sufficient data are available within each to allow estimation of a different selectivity pattern.

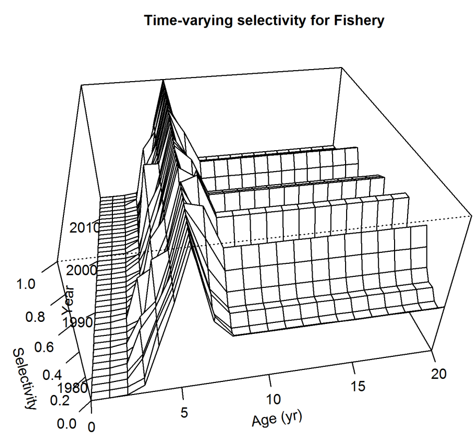
A new procedure for accounting for fishery selectivity has been proposed here in model 15.6 where an age-varying random walk is used to characterize the selectivity for all combined fisheries (trawl, longline and pots) each year. This procedure seems attractive given the high level of partitioning required for model 11.5. If such a procedure can provide a means of accounting for total fishery removals each year according to size/age, then it should be acceptable. Diagnostic plots for fishery lengths, both by year and combined for model 15.6, show rather good fits to available data (all residuals are also within the range -2.0 to 2.0). There is very little catch taken aged above about 8, so fixing selectivity above that age seems reasonable.

As the proportion of trawl catch to longline has changed considerably over time, it would be expected that large changes in the general pattern of selectivity would also be observed, that are somewhat evident in the plot (Figure 9), but of possible concern. Is the amount of change consistent with the broad movement of the fishery from trawl to longline over time?

Also of some concern is that the general fishery pattern for model 15.6 is dome-shaped, allowing the model some flexibility to generate cryptic spawning biomass. This is also an area of on-going work, and some diagnostics associated with it are in development or available from Github as additions to R4SS. At present, the available code only works for 2 sex models, so cannot be applied here, but could be further generalized to do so. The inclusion of surveys that are more directed towards the older fish in the population help to alleviate cryptic biomass problems, and is therefore a further reason to consider the addition of at least one longline survey to the base model.

I believe that options are only currently available in SS for a random walk by age for annual selectivity, as used for model 15.6. If the same was done by length, more parameters would be required (if 1cm size bins), or alternative bin patterns could be explored. Such a length-based exploration would be useful, should such capability be available in SS.

Figure 9 EBS time-varying age-based fishery selectivity from model 15.6.



As many current SS assessments grapple with highly partitioned fishery data, such a procedure has the potential for resolving some of those problems also. I do not have previous personal experience with this procedure, and am reluctant to agree on its use without a supporting simulation study that confirms its equivalence or even superiority to a high degree of data partitioning. Such a study would be reasonably easy to design and carry out. However, I am willing to agree that it seems to provide a good resolution to the problem for the fishery selectivity in the EBS models.

2.2.2d *Should the models be structured with respect to gear type?*

As this question mostly relates to dealing with the fisheries and not surveys, the discussion under ToR 2.2.2c was generalized to address both season and gear type.

2.2.2e *How much time variability should be allowed, and in which parameters?*

The only population biological parameter allowed to vary with time in most SS stock assessments is annual recruitment levels. Cumulative information on annual recruitment strength is provided fairly directly by composition data, so the reasons especially for high peaks and troughs in recruitment are usually apparent in the available data. It has also been recognized that other parameters are likely to vary through time – in particular natural mortality, but also growth and maturity. For natural mortality it has been considered difficult to estimate time trends in changes without strong independent estimates for those changes, such as from ecosystem studies showing differences in predator abundance, and that time trends in M are difficult to disentangle from other factors such as catch mis-specification (e.g. see Brodziak et al., 2011). Allowing time variation in factors that directly affect productivity also lead to questions about choice of appropriate time periods for the selection of management reference points, and how to make appropriate stock projections.

Additional model parameters that may vary with time that are often dealt with using time-block methods are fishery/survey selectivity and catchability. As already mentioned, for fisheries that are not associated with an abundance index, a fairly freely estimated time-varying pattern (such as used for EBS model 15.6) may be acceptable if it suitably captures annual fishery removals by size/age. For surveys the situation differs. Surveys are the most important source of abundance information for the model, particularly because at least the gear selectivity can be maintained as a constant through time. Availability (either by age or year) is another matter, but is usually treated as a source of additional random error. If a true trend (or even a step) exists in either survey selectivity or catchability, then that survey is biased, and the bias needs to be accounted for, or the survey truncated, split or discarded. Such a bias would ideally be investigated and identified with a focused study and auxillary data not necessarily used in the assessment model. Adding annual time-variability to survey selectivity or catchability and finding that trends are estimated may simply be providing a means for the model to trade trends in population abundance to improve the fit to noisy composition data in preference to abundance indices. The reason that such a model might result in trends in survey selectivity or catchability are not readily apparent from standard input data sources, and may be difficult to diagnose. Results from estimation of annual variability for the EBS trawl survey catchability in model 15.6 (Figure 10) do exhibit some runs in residuals that may be of concern – particularly from 1993 to 1996. Time-changes in trawl survey selectivity as estimated by the EBS model 15.6 shows very little change through time, suggesting that time-variability in trawl survey selectivity as implemented is not required (Figure 11).

My own recommendation for now is that time variability should be allowed in a parameter when there is an available reliable data source that fairly directly measures such a change, and that a trend exists in that data source that needs to be captured by the assessment model. This situation only currently exists for recruitment and fishery selectivity in the EBS model. It also provides some support to consider time variability in weight-at-length or size-at-age if those data sets show considerable trends over time.

Others (e.g. Anders Nielsen, Jim Thorson) have proposed that a more appropriate way to deal with time variability is to use mixed-effects models with time-varying “nuisance” variables such as recruitment modeled as random effects. Improved solutions for time-varying parameters may be possible using all of the currently available data sources, if/when SS RE becomes available.

Figure 10 EBS time-varying trawl survey catchability from model 15.6.

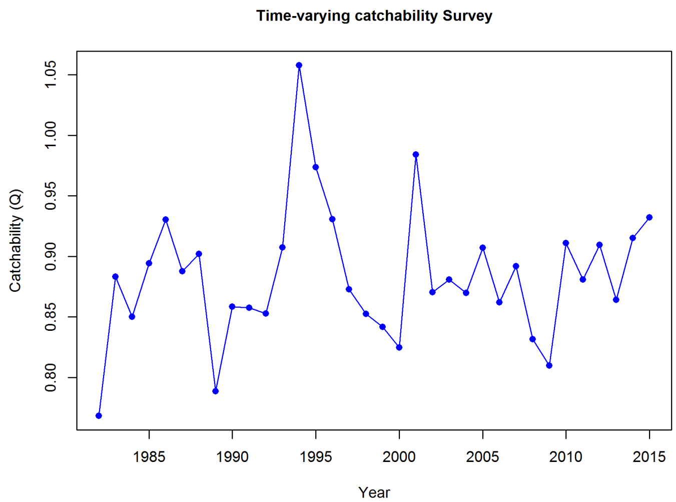
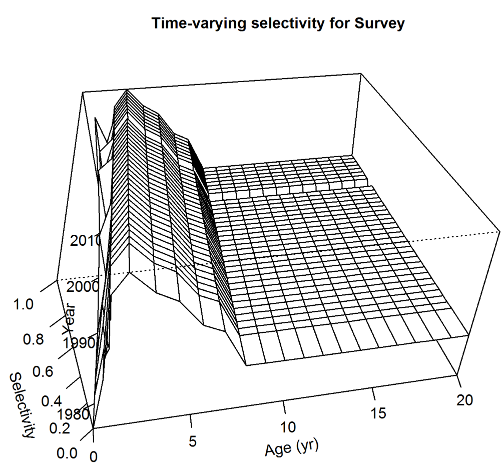


Figure 11 EBS time-varying trawl survey selectivity from model 15.6.



2.2.2f *What constraints, if any, should be placed on survey selectivity at older ages?*

The models examined during the review for the EBS seem to fairly clearly demonstrate that the trawl survey selectivity is dome-shaped. However, the possibility that the survey is in fact asymptotic has not been eliminated. The extent of the survey dome-shape may, for example, be confounded with M. It may be that different data sources are in conflict about the estimated value for M that can be diagnosed with a Piner profile plot of likelihood components. Exploration of age-specific M (e.g. starting with a Lorenzen function) could also be done.

A range of plausible alternative models should be explored, and the extent of the estimated dome selectivity for the trawl survey examined for each to see if the dome is consistently required. However, as the extent of the trawl survey dome is probably one of the major axes of uncertainty in the model at present, it should remain freely estimated and informed by the available data in any chosen base model, possibly with forcing more or less dome as sensitivity analyses in the final assessment.

2.2.2g *What constraints, if any, should be placed on survey catchability?*

Because of the history of the development and use of the trawl survey as an absolute index of abundance, there remains some belief that there is sufficient information available to determine at least a plausible acceptable range for survey q, and to some, that range could be perceived to be quite narrow. Much work has been directed towards net avoidance and how that might be compensated by a q adjustment. I believe that all major potential sources of error in survey q should at least be stated in an accessible document, and errors in those dimensions at least be qualitatively examined and ranked. Those should include avoidance and other gear-specific fish behavioral issues, and potential error in scaling the swept area estimates to the population using assumptions about the population distribution during the survey by depth and area, and also even the assumption of known stock boundaries. A qualitative evaluation such as this would probably make it clear that the true error in q is reasonably high. It would also assist to determine what priorities should be given to field studies that may be directed towards reduction of the error in survey q and adjustments required to scale area swept biomass estimates to the total (available given survey selectivity) population. An extension to a more quantitative evaluation of the potential errors may also lead to a prior distribution for EBS shelf bottom trawl survey q that can be generally agreed, and could then be used for modeling without much controversy. Without at least a comprehensive qualitative evaluation of all major error sources, decisions about rejection of models that estimate q based on how different the estimated q is from acceptable values remains difficult, and currently in the domain of pragmatic judgment.

I believe that models that estimate the shelf bottom trawl survey q using a fairly non-informative prior (as in model 15.6) should currently be preferred. Agreed bounds on prior survey q point estimates can be used as one of the acceptance criteria for particular models. I personally have a fairly high tolerance for those values (based however, on only a limited background knowledge for this particular survey), and am comfortable with at least a factor of 2.0 (0.5 – 2.0 times the initial point estimates).

Should additional surveys be added to the models, q values for Pacific cod for those are less well understood, and non-restrictive priors for those are preferable, with q estimated.

2.2.2h *How should large gradients be dealt with in otherwise apparently converged models?*

Large gradients are generally considered to be an indication of a problem. However, if the hessian can be inverted and jitters also indicate convergence, then perhaps the problem is only minor. I do not have any reason to doubt the explanation given in the EBS assessment document for why large gradients might occur, but it does suggest to me that the implementation of age selectivity pattern 17 requires a closer look to determine if the problem can be corrected (e.g. to determine whether it contains badly behaved/non-differentiable “if” statements).

2.2.2i ***Anything else on which the reviewers care to comment.***

Retrospectives

Diagnosis of retrospective bias in stock assessments has received considerable past attention in the literature and was also the subject of a BSAI/GOA working group in 2013 according to meeting background information. Despite this attention, research is on-going, and means for diagnosis and correction for retrospective patterns are not agreed. Several diagnostic measures are available including Mohn's ρ, the so-called Woods Hole ρ, and the RMSE method devised by the BSAI/GOA working group. I am familiar with two rules of thumb that can be used to diagnose retrospective patterns that need to be addressed in some way. The first and simplest is by Hurtado-Ferro et al. (2014) that says that "values of Mohn's ρ higher than 0.20 or lower than -0.15 for longer-lived species (upper and lower bounds of the 90% simulation intervals for the flatfish base case), or higher than 0.30 or lower than -0.22 for shorter-lived species (upper and lower bounds of the 90% simulation intervals for the sardine base case) should be cause for concern and taken as indicators of retrospective patterns." The second by Brooks and Legault (2015) from VPA assessments "is to plot the terminal year estimate of SSB(T) vs F(T) along with bootstrap percentiles and compare that to the point estimate when SSB(T) and F(T) are adjusted by $\rho_{SSB,7}$ and $\rho_{F,7}$, respectively" to see if the ρ-adjusted point estimate falls outside the bootstrap percentiles on either axis - see Brooks and Legault (2015) for details. Brooks and Legault (2015) also provide a procedure for adjustment of short-term projection results to account for substantial retrospective patterns. Ideally, the diagnostics for a model acceptable for use for management advice should not show significant retrospective bias. EBS model 11.5 and the initial AI SS models did show significant retrospective bias (at least according to the Hurtado-Ferro et al. (2014) rule of thumb) that indicated that results from those models are not reliable for use for management advice, and that improved alternative models should be sought, or at least a projection correction may be required. Further model explorations for both regions have found models that do not exhibit a strong retrospective bias, and on that basis would be judged as improved models. Retrospective bias provides evidence for model mis-specification, but of course, the lack of a retrospective bias does not prove that the model is correctly specified.

So-called Ianelli "squid plots" provide an additional useful means for looking at retrospective patterns in annual recruitment deviations, but have potential application to any parameter allowed to deviate annually in a model.

Catch uncertainty

As for many models, historical catch in particular is uncertain, and the best estimate of historical catch has been made using assumptions that seem supportable. However, the construction of alternative plausible historical catch scenarios would be useful for the determination of sensitivity of the model to that uncertainty.

Steepness

Tier 3 methods by default assume a steepness value of 1.0. A requested run using a steepness value of 0.7 shows that EBS results are somewhat sensitive to the choice of steepness value, and this dimension of uncertainty should be highlighted.

Regime change

A regime change in 1976-77 affecting log mean recruitment in EBS model 11.5 has been avoided in EBS model 15.6 by starting the latter model after the regime change. Shifts in 1989 and 1999 have also been suggested according to the ecosystem considerations in the assessment documentation. Regime change was not examined at all during the review, but is another potential source of model uncertainty.

Inclusion of marginal age composition vs CAAL data

At present, both the EBS and AI enter age-at-length data as marginal age distributions. There has been a gradual trend in stock assessments to make improved use of data from otoliths by entering the data into models as conditional age-at-length. During the review the general wisdom of this approach was questioned as it was mentioned that some recent assessments had reverted back to marginal age distributions. A standard approach for dealing with age-at-length data currently seems to be unavailable.

References

Brodziak, J., Ianelli, J., Lorenzen, K. and Methot, R.D. 2011. Estimating Natural Mortality in Stock Assessment Applications. NOAA Technical Memorandum NMFS-F/SPO-199.

Brooks, E.N. and Legault, C.M. 2015. Retrospective forecasting – evaluating performance of stock projections. Canadian Journal of Fisheries and Aquatic Sciences. doi: 10.1139/cjfas-2015-0163.

Hurtado-Ferro, F., Szuwalski, C. S., Valero, J. L., Anderson, S. C., Cunningham, C. J., Johnson, K. F., Licandeo, R., McGilliard, C. R., Monnahan, C. C., Muradian, M. L., Ono, K., Vert-Pre, K. A., Whitten, A. R., and Punt, A. E. 2014. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. ICES J. Mar. Sci. doi: 10.1093/icesjms/fsu198.

Punt, A. E. (*In Press*). Some insights into data weighting in integrated stock assessments. Fisheries Research.

Thorson, J.T., 2014. Standardizing compositional data for stock assessment. ICES J. Mar. Sci. J. Cons. 71, 1117–1128.

Appendix 1: Bibliography of materials provided for review

Assessment of the Pacific cod stock in the eastern Bering Sea (220 p.), including a history of alternative models developed for assessing Pacific cod in the EBS (Appendix 2.3)

Assessment of the Pacific cod stock in the Aleutian Islands (143 p.), including a history of alternative models developed for assessing Pacific cod in the AI (Appendix 2A.3)

Comments on the final 2015 EBS and AI Pacific cod assessments by the Plan Team and SSC

Additional documents

Excerpt from the BSAI Groundfish Plan Team minutes of November 2015

Excerpt from the SSC minutes of December 2015

FLC comments

Thompson, G. Specifying the standard deviations of randomly time-varying parameters in stock assessment models based on penalized likelihood: a review of some theory and methods.

Weinberg et al. Is Pacific cod (*Gadus macrocephalus*) survey selectivity dome-shaped? Direct evidence from trawl studies.

Appendix 2: A copy of the CIE Statement of Work

Statement of Work

External Independent Peer Review by the Center for Independent Experts

Assessment of the Pacific cod stocks in the Eastern Bering Sea and Aleutian Islands

Scope of Work and CIE Process: The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Technical Representative (COTR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide impartial and independent peer review without conflicts of interest. CIE reviewers are selected by the CIE Steering Committee and CIE Coordination Team to conduct the independent peer review of NMFS science in compliance the predetermined Terms of Reference (ToRs) of the peer review. Each CIE reviewer is contracted to deliver an independent peer review report to be approved by the CIE Steering Committee and the report is to be formatted with content requirements as specified in **Annex 1**. This SoW describes the work tasks and deliverables of the CIE reviewer for conducting an independent peer review of the following NMFS project. Further information on the CIE process can be obtained from www.ciereviews.org.

Project Description: Despite exploration of a large number of alternative models and multiple levels of review each year, the annual assessments of the Pacific cod stocks in the EBS and AI continue to be controversial. Of particular concern currently is the estimation of catchability and selectivity for the bottom trawl survey in each area. However, review is requested of all aspects of the stock assessment models. The combined Pacific cod fisheries in the EBS and AI are of great economic importance, ranking second only to pollock in recent years. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements for CIE Reviewers: Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. CIE reviewers shall have working knowledge and recent experience in the application of stock assessment methods in general, and preferably Stock Synthesis in particular. Each CIE reviewer's duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting ***scheduled in Seattle, WA during February 16-19, 2016.***

Statement of Tasks: Each CIE reviewers shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Prior to the Peer Review: Upon completion of the CIE reviewer selection by the CIE Steering Committee, the CIE shall provide the CIE reviewer information (full name, title, affiliation, country, address, email) to the COTR, who forwards this information to the NMFS Project Contact no later the date specified in the Schedule of Milestones and Deliverables. The CIE is responsible for providing the SoW and ToRs to the CIE reviewers. The NMFS Project Contact is responsible for providing the CIE reviewers with the background documents, reports, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact is also responsible for providing the Chair a copy of the SoW in advance of the panel review meeting. Any changes to the SoW or ToRs must be made through the COTR prior to the commencement of the peer review.

Foreign National Security Clearance: When CIE reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for CIE reviewers who are non-US citizens. For this reason, the CIE reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/ http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE Lead Coordinator on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

Assessment of the Pacific cod stock in the eastern Bering Sea (220 p.), including a history of alternative models developed for assessing Pacific cod in the EBS (Appendix 2.3)

Assessment of the Pacific cod stock in the Aleutian Islands (143 p.), including a history of alternative models developed for assessing Pacific cod in the AI (Appendix 2A.3)

Comments on the final 2015 EBS and AI Pacific cod assessments by the Plan Team and SSC

Panel Review Meeting: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs cannot be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COTR and CIE Lead Coordinator.** Each CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein. The NMFS Project

Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements).
The NMFS Project Contact is responsible for ensuring that the Chair understands the contractual role of the CIE reviewers as specified herein. The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

The review meeting will include three main parts: The first will consist of a series of presentations with follow-up questions and discussions by CIE reviewers, and will be chaired by an AFSC scientist or supervisor. The second will consist of real-time model runs and evaluations conducted in an informal workshop setting, and will be chaired jointly by the CIE reviewers. The third, time permitting, will consist of initial report writing by the CIE reviewers, with opportunity for additional questions of the assessment author.

<u>Contract Deliverables - Independent CIE Peer Review Reports</u>: Each CIE reviewer shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

<u>Other Tasks – Contribution to Summary Report</u>: Each CIE reviewer may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. Each CIE reviewer is not required to reach a consensus, and should provide a brief summary of the reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

Specific Tasks for CIE Reviewers: The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate during the panel review meeting ***scheduled at the Alaska Fisheries Science Center in Seattle, WA during February 16-19, 2016.***
3) Participate at the peer review meeting ***tentatively scheduled at the Alaska Fisheries Science Center in Seattle, WA during February 16-19, 2016*** as specified herein, and conduct an independent peer review in accordance with the ToRs (**Annex 2**).
4) No later than ***March 4, 2016***, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Dr. Manoj Shivlani, CIE Lead Coordinator, via email to mshivlani@ntvifederal.net, and CIE Regional Coordinator, via email to Dr. David Die ddie@rsmas.miami.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

Schedule of Milestones and Deliverables: CIE shall complete the tasks and deliverables described in this SoW in accordance with the following ***tentative*** schedule.

January 11, 2016	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
February 1, 2016	NMFS Project Contact sends the CIE Reviewers the pre-review documents
February 16-19, 2016	Each reviewer participates and conducts an independent peer review during the panel review meeting
March 4, 2016	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
March 18, 2016	CIE submits CIE independent peer review reports to the COTR
March 25, 2016	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: This 'Time and Materials' task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery Management Council, and Council's SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COTR within 10 working days after receipt of all required information of the decision on changes. The COTR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

Acceptance of Deliverables: Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COTR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COTR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COTR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:
(1) The CIE report shall completed with the format and content in accordance with **Annex 1**, (2) The CIE report shall address each ToR as specified in **Annex 2**,
(3) The CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

Distribution of Approved Deliverables: Upon acceptance by the COTR, the CIE Lead Coordinator shall send via e-mail the final CIE reports in *.PDF format to the COTR. The COTR will distribute the CIE reports to the NMFS Project Contact and Center Director.

Support Personnel:

Allen Shimada
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
Allen Shimada@noaa.gov Phone: 301-427-8174

Manoj Shivlani, CIE Lead Coordinator
Northern Taiga Ventures, Inc.
10600 SW 131st Court, Miami, FL 33186
mshivlani@ntvifederal.com Phone: 305-968-7136

Key Personnel:

NMFS Project Contact:

Grant Thompson, Alaska Fisheries Science Center c/o
Department of Fisheries and Wildlife
Oregon State University
104 Nash Hall, Corvallis, OR
97331 Phone: 541-737-9318
grant.thompson@noaa.gov

Annex 1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including providing a brief summary of findings, of the science, conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the Summary Report that they feel might require further clarification.

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The CIE independent report shall be an independent peer review of each ToRs, and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Statement of Work
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Annex 2: Terms of Reference for the Peer Review

Assessment of the Pacific cod stocks in the Eastern Bering Sea and Aleutian Islands

1. Evaluate and provide recommendations on data used in the assessment models. In particular:
    a. Should data from the IPHC longline survey be used in either assessment?
    b. Should data from the NMFS longline survey be used in either assessment?
2. Evaluate and provide recommendations on model structure, assumptions, and estimation procedures. In particular:
    a. How should the various data sets be weighted?
    b. What form (i.e., Stock Synthesis "pattern") should be used for the selectivity functions?
    c. Should the models be structured with respect to season?
    d. Should the models be structured with respect to gear type?
    e. How much time variability should be allowed, and in which parameters?
    f. What constraints, if any, should be placed on survey selectivity at older ages?
    g. What constraints, if any, should be placed on survey catchability?
    h. How should large gradients be dealt with in otherwise apparently converged models?
    i. Anything else on which the reviewers care to comment.

Annex 3: Tentative Agenda

CIE Review of the EBS and AI Pacific cod stock assessment models

Alaska Fisheries Science Center
7600 Sand Point Way NE, Seattle, WA
98115 February 16-19, 2016
Building 4; Room 2039 (except Wednesday afternoon), Room 2143 (Wednesday afternoon)

Review panel chair: Anne Hollowed, Anne.Hollowed@noaa.gov
Senior assessment author: Grant Thompson, Grant.Thompson@noaa.gov
Security and check-in: Sandra Lowe, Sandra.Lowe@noaa.gov (206)526-4230

Sessions will run from 9 a.m. to 5 p.m. each day, with time for lunch and morning and afternoon breaks. Discussion will be open to everyone, with priority given to the panel and senior assessment author.

Tuesday, February 16

Preliminaries:
- 09:00 Introductions and adoption of agenda—Anne

Data sources (current and potential):
- 09:10 Overview of data types used in the assessments—Grant
- 09:20 Catch accounting system and in-season management—AKRO SF Division (via WebEx)
- 09:50 Observer program—AFSC FMA Division
- 10:20 Break
- 10:30 EBS trawl survey—AFSC RACE Division
- 11:00 AI trawl survey—AFSC RACE Division 11:30 IPHC longline survey—IPHC
- 12:00 Lunch
- 13:00 NMFS longline survey—AFSC Auke Bay Laboratory (via WebEx)
- 13:30 Age composition and mean-length-at-age data—AFSC REFM Division

Assessment models:
- 14:00 Assessment history—Grant
- 15:00 Break
- 15:10 Current assessments—Grant 16:10 Discussion—Everyone
- 16:40 Assignments for models to be presented on Wednesday—Panel

Wednesday, February 17 and Thursday, February 18

Review of models assigned the previous day—Grant
Discussion, real-time model runs—Everyone
Assignments for models to be presented the following day—Panel

Friday, February 19

Review of models assigned on Thursday—Grant
Discussion, real-time model runs—Everyone
Report writing (time permitting)—Panel

Appendix 3: List of participants

Participants List
CIE Review of the EBS and AI Pacific cod stock assessment models
Alaska Fisheries Science Center (AFSC)
7600 Sand Point Way NE, Seattle, WA 98115
February 16-19, 2016

Review panel chair: Anne Hollowed, Anne.Hollowed@noaa.gov
Senior assessment author: Grant Thompson, Grant.Thompson@noaa.gov
Support assessment author: Steve Barbeaux, Steve.Barbeaux@noaa.gov

Name	Organization	e-mail
Craig Faunce	AFSC	Craig.Faunce@noaa.gov
Craig Kastell	AFSC	Craig.Kastell@noaa.gov
Delsa Anderl	AFSC	Delsa.Anderl@noaa.gov
Liz Chilton	AFSC	Elizabeth.Chilton@noaa.gov
Ron Felthoven	AFSC	Ron.Felthoven@noaa.gov
Sandi.Neidetcher	AFSC	Sandi.Neidetcher@noaa.gov
Thomas Helser	AFSC	Thomas.Helser@noaa.gov
Tom Wilderbuer	AFSC	Tom.wilderbuer@noaa.gov
Farron Wallace	AFSC – Chair NPFMC SSC	Farron.Wallace@noaa.gov
Jeff Napp	AFSC – RACE Division Leader	Jeff.Napp@noaa.gov
Dana Hanselman	AFSC- Auke Bay Laboratory	Dana.Hanselman@noaa.gov
Laura Junge	Alaska Pacific University	Laura.junge@gmail.com
Anne Vanderhoeven	Bristol Bay Economic Development Corporation	Anne.bbedc@ak.net
Jean-Jacques Maguire	Center of Independent Experts (CIE)	jeanjacquesmaguire@gmail.com
Neil Klaer	CIE	Neil.Klaer@luggara.com
Robin Cook	CIE	Melford@clara.co.uk
Chad See	Freezer Longline Coalition	Chadsee@freezerlongline.bz
Gerry Merrigan	Freezer Longline Coalition	Merrigan@oci.net
Mike Peterson	Interested Fisherman	
Patrick Lynch	NMFS Science and Technology	Patrick.Lynch@noaa.gov
Mary Furuness	NMFS-Alaska Regional Office	Mary.Furuness@noaa.gov
Diana Stram	North Pacific Fishery Management Council (NPFMC)	Diana.Stram@noaa.gov
Gwladys Lambert	NRC post-doc	Gwladys.Lambert@noaa.gov
Jennifer Cahalan	Pacific States Marine Fish Commission/AFSC	Jcahalan@pfmfc.gov
Steve Martell	SeaState	Steve@seastateinc.com